Reading: VDV Chapters 18 and 19; Notes on the Arzelà-Ascoli theorem on the course website.

1 Recap: Uniform Limits in Distributions

Definition 1.1. A process \((X_n)_{n=1}^{\infty}\), \(X_n \in L^\infty(T)\), is asymptotically stochastically equi-continuous (ASEC) if \(\forall \varepsilon > 0, \eta > 0\), there exists a partition \(T_1, \ldots, T_m\) of \(T\) such that
\[
\limsup_{n \to \infty} P\left( \max_{1 \leq i \leq m} \sup_{s,t \in T_i} |X_n(s) - X_n(t)| \geq \varepsilon \right) \leq \eta. \tag{1}
\]

2 Weak Convergence in \(L^\infty(T)\)

Theorem 1. Let \(\{X_n\}_{n=1}^{\infty} \subset L^\infty(T)\) be a sequence of stochastic processes on \(T\). The followings are equivalent.

1. \(X_n\) converge in distribution to a tight stochastic process \(X \in L^\infty(T)\);
2. both of the followings:
   a. Finite Dimensional Convergence (FIDI): for every \(k \in \mathbb{N}\) and \(t_1, \ldots, t_k \in T\),
      \(\big(X_n(t_1), \ldots, X_n(t_k)\big)\)
      converge in distribution as \(n \to \infty\);
   b. the sequence \(\{X_n\}\) is asymptotically stochastically equicontinuous.

Proof (1) \(\Rightarrow\) (2) is trivial. Here we only prove (2) \(\Rightarrow\) (1).

Part I: Consider countable subsets of \(T\).
Let \(m \in \mathbb{N}\), and construct partitions \(T_1^m, \ldots, T_k^m\) of \(T\) such that
\[
\limsup_{n \to \infty} P\left( \max_{1 \leq i \leq k} \sup_{s,t \in T_i^m} |X_n(s) - X_n(t)| \geq 2^{-m} \right) \leq 2^{-m}. \tag{2}
\]

Without loss of generality, assume that \(\{T_i^m\}_m\) are nested partitions. For each \(m \in \mathbb{N}\), define
\[
\rho_m(s, t) = \begin{cases} 0 & \text{if } s, t \in T_i^m \text{ for some } i \\ 1 & \text{otherwise} \end{cases}
\]
and let
\[
\rho(s, t) = \sum_{m=1}^{\infty} 2^{-m} \rho_m(s, t) \quad \forall s, t \in T.
\]
It is easy to see that $\rho$ is a metric. Also notice that if $s, t \in T_i^m$, there is $\rho(s, t) < 2^{-m}$. Therefore $\text{diam}(T_i^m) < 2^{-m}$. From each $T_i^m$ we pick out an element $t_{i,m}$ and define

$$T_0 = \bigcup_{m=1}^{\infty} \bigcup_{i=1}^{k_m} \{t_{i,m}\}.$$  

Obviously $T_0$ is countable. Further, for each $t \in T$ and any $m \in \mathbb{N}$, suppose $t \in T_j^m$, we have $\rho(t, t_{j,m}) < 2^{-m}$. Hence $T_0$ is dense in $T$ with respect to the $\rho$-metric.

**Part II: Use $T_0$ to obtain a limit process in $C(T, \mathbb{R})$.**

By Kolmogorov’s extension theorem, there is some stochastic process $\{X(t)\}_{t \in T_0}$ where

$$(X_n(t_1), \ldots, X_n(t_k)) \xrightarrow{d} (X(t_1), \ldots, X(t_k)) \quad \forall k \in \mathbb{N}, \ t_1, \ldots, t_k \in T_0. \quad (3)$$

Let $S$ be a finite subset of $T_0$, then

$$\mathbb{P} \left( \sup_{s,t \in T_0, \ \rho(s,t) < 2^{-m}} |X(s) - X(t)| \geq 2^{-m} \right) \leq \mathbb{P} \left( \max_{1 \leq i \leq k_m} \sup_{s,t \in T_i^m \cap T_0} |X(s) - X(t)| \geq 2^{-m} \right) \leq \lim_{S \uparrow T_0} \mathbb{P} \left( \max_{1 \leq i \leq k_m} \max_{s,t \in T_i^m \cap S} |X(s) - X(t)| \geq 2^{-m} \right) \leq \lim_{S \uparrow T_0} \limsup_{n \to \infty} \mathbb{P} \left( \max_{1 \leq i \leq k_m} \max_{s,t \in T_i^m \cap S} |X_n(s) - X_n(t)| \geq 2^{-m} \right) \leq 2^{-m}, \quad (4)$$

where (a) is because $\rho(s,t) < 2^{-m}$ implies $s,t \in T_i^m$ for some $i \leq k_m$; (b) follows from monotone convergence theorem; (c) is the result of finite dimensional convergence (FIDI); (d) results from (2). Notice that

$$\sum_{m=1}^{\infty} \mathbb{P} \left( \sup_{s,t \in T_0, \ \rho(s,t) < 2^{-m}} |X(s) - X(t)| \geq 2^{-m} \right) \leq \sum_{m=1}^{\infty} 2^{-m} = 1,$$

By Borel-Cantelli lemma, we have

$$\mathbb{P} \left( \exists M \in \mathbb{N}, \ s.t. \ \forall m \geq M, \ \sup_{s,t \in T_0, \ \rho(s,t) < 2^{-m}} |X(s) - X(t)| < 2^{-m} \right) = 1.$$  

Therefore with probability 1, process $\{X(t)\}_{t \in T_0}$ is continuous (even locally Lipschitz), i.e. $X \in C(T_0, \mathbb{R})$. Since $T_0$ is dense in $T$, we have that

$$X \in C(T, \mathbb{R}) \quad \text{a.s.}$$

Also notice that the total boundedness of $T$ implies the uniform continuity of $X$.

**Part III: Show that $X_n \xrightarrow{d} X$.**

We only have to show $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for all bounded and Lipschitz $f$. Recall that $t_{i,m}$ is an element of $T_i^m$. For any $t \in T_i^m$, let $\pi_m(t) = t_{i,m}$. Then there is $\rho(t, \pi_m(t)) < 2^{-m}$. Define $(X \circ \pi_m)(t) = X(\pi_m(t))$, then we have

$$X \circ \pi_m \xrightarrow{a.s.} X, \quad \text{as} \ m \to \infty,$$
by uniform continuity. In other words,
\[
\sup_{t \in T} |(X \circ \pi_m)(t) - X(t)| \to 0, \quad \text{as } m \to \infty.
\] (5)

Using finite dimensional converge, there is also
\[
X_n \circ \pi_m \overset{d}{\to} X \circ \pi_m, \quad \text{as } n \to \infty.
\] (6)

For \( f : L^\infty(T) \mapsto [0,1] \) that is Lipschitz, by triangular inequality,
\[
|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| \leq |\mathbb{E}[f(X_n \circ \pi_m)] - \mathbb{E}[f(X)]| + |\mathbb{E}[f(X \circ \pi_m)] - \mathbb{E}[f(X)]|.
\] (7)

Notice that from (6) we have
\[
|\mathbb{E}[f(X_n \circ \pi_m)] - \mathbb{E}[f(X \circ \pi_m)]| \to 0, \quad \text{as } n \to \infty, \quad \forall m.
\] (8)

From (5) together with the boundedness of \( f \), there is
\[
|\mathbb{E}[f(X \circ \pi_m)] - \mathbb{E}[f(X)]| \to 0, \quad \text{as } m \to \infty.
\] (9)

Finally we also have
\[
|\mathbb{E}[f(X_n \circ \pi_m)] - \mathbb{E}[f(X_n)]| \stackrel{(e)}{\leq} \|f\|_{\text{Lip}} \cdot \mathbb{E}\left[1 \land \|X_n \circ \pi_m - X_n\|_{\infty}\right]
\leq \|f\|_{\text{Lip}} \cdot \left(\epsilon + \mathbb{P}\left(\sup_{t \in T} |X_n(t) - X_n(\pi_m(t))| \geq \epsilon\right)\right),
\]
where \( \epsilon > 0 \) is arbitrary, \( \|f\|_{\text{Lip}} \) is the Lipschitz constant of \( f \), and \( (e) \) originates from the Lipschitz and boundedness of \( f \). Setting \( \epsilon = 2^{-m} \) and taking \( n \to \infty \), there is
\[
\limsup_{n \to \infty} |\mathbb{E}[f(X_n \circ \pi_m)] - \mathbb{E}[f(X_n)]| \leq \|f\|_{\text{Lip}} \cdot \limsup_{n \to \infty} \left(2^{-m} + \mathbb{P}\left(\sup_{t \in T} |X_n(t) - X_n(\pi_m(t))| \geq 2^{-m}\right)\right)
\leq \|f\|_{\text{Lip}} \cdot (2^{-m} + 2^{-m}),
\]
where \( (f) \) is the result of the asymptotic stochastic equicontinuity of \( \{X_n\} \). Combining (7), (8), (9) and (10), the proof is complete. \( \square \)

**Remark** We actually showed that the limit process \( (X_t)_{t \in T} \) has uniformly continuous sample paths for some metric \( \rho \) with probability 1, where \( (T, \rho) \) is totally bounded.

**Corollary 2.** Suppose that \( (T, d) \) is a totally bounded metric space with
\[
\limsup_{n \to \infty} \mathbb{P}\left(\sup_{d(s,t) < \delta} |X_n(s) - X_n(t)| \geq \varepsilon\right) = 0,
\] (11)

and has FIDI, then \( X_n \overset{d}{\to} X \in L^\infty(T) \), \( X \) is continuous w.p.1.

**Proof** Show ASEC: for \( \varepsilon > 0, \delta > 0 \), choose a partition of \( T \), \( \{T_i\}_{i=1}^m \), with \( \text{diam}(T_i) < \delta \), then
\[
\max_{i} \sup_{(s,t) \in T_i} |X_n(s) - X_n(t)| \leq \sup_{d(s,t) < \delta} |X_n(s) - X_n(t)|.
\] (12)

The proof is complete. \( \square \)
3 Donsker Classes

Definition 3.1. A collection $\mathcal{F}$ of functions is called $P$-Donsker if the process

$$\left(\sqrt{n}(P_n - P)f\right)_{f \in \mathcal{F}}$$

converges to a tight limit process in $L^\infty(\mathcal{F})$, i.e. $\sqrt{n}(P_n - P)$ converges in $L^\infty(\mathcal{F})$.

Remark This limit process must be a Gaussian process $\mathbb{G} = \mathbb{G}_P$, i.e. $\mathbb{G}$ is a random mapping from $\mathcal{F}$ to $\mathbb{R}$ such that

$$(\mathbb{G}f_1, \cdots, \mathbb{G}f_k) \sim N\left(0, \left[Cov_{P}(f_i,f_j)\right]_{i,j=1}^{k}\right)\quad \forall f_1, \cdots, f_k \in \mathcal{F}, \ k < \infty,$$

where

$$Cov_{P}(f_i,f_j) = Cov_{X \sim P}[f_i(X), f_j(X)].$$

Example 1: ($P$-Brownian bridge) Let $F_n(t) = P_n(X \leq t)$, $F(t) = P(X \leq t)$, and $\mathcal{F} = \{1(\cdot \leq t)\}_{t \in \mathbb{R}}$. Then

$$\left\{\sqrt{n}(F_n(t) - F(t))\right\}_{t \in \mathbb{R}} \stackrel{d}{\to} \mathbb{G}_P \subset L^\infty(\mathbb{R}).$$

For $s,t \in \mathbb{R}$,

$$E[1(X \leq s)1(X \leq t)] = F(s \wedge t),$$

then $\mathbb{G}$ is a Gaussian process with

$$Cov(\mathbb{G}_t, \mathbb{G}_s) = F(s \wedge t) - F(s)F(t),$$

and $\mathbb{G}_t - \mathbb{G}_s$ is Gaussian, and

$$Var(\mathbb{G}_t - \mathbb{G}_s) = E\left[G^2 + G_t^2\right] - 2E[G_sG_t] = F(s)(1 - F(s)) + F(t)(1 - F(t)) - 2F(s \wedge t) + 2F(s)F(t).$$

Example 2: (Lipschitz functions) Let $\Theta \subset \mathbb{R}^d$, where $\Theta$ is compact. Let $\ell : \Theta \times \mathcal{X} \mapsto \mathbb{R}$, with $\ell(\cdot, x)$ is $L(x)$--Lipschitz on $\Theta$, and $E_P[L(x)^2] < \infty$, then $\mathcal{F} = \{\ell(\theta, \cdot)\}_{\theta \in \Theta}$ is $p$–Donsker, and

$$\sqrt{n}(P_n\ell(\cdot, x) - P\ell(\cdot, x)) \stackrel{d}{\to} \mathbb{G} \subset C(\Theta, \mathbb{R}),$$

with

$$Cov(\mathbb{G}_{\theta_0} - \mathbb{G}_{\theta_1}) = Cov(\ell(\theta_0, x), \ell(\theta_1, x)).$$

The following theorem shows that, a function class is $P$-Donsker if it has uniformly bounded entropy.
Theorem 3. Let \( \mathcal{F} \) be a class of functions mapping \( \mathcal{X} \) to \( \mathbb{R} \), and \( F : \mathcal{X} \to \mathbb{R} \) be an envelope of \( \mathcal{F} \), i.e.

\[
f \in \mathcal{F} \Rightarrow |f(x)| \leq |F(x)|, \forall x \in \mathcal{X}.
\]

Suppose that

\[
\int_{0}^{\infty} \sup_{Q} \sqrt{\log N(\mathcal{F}, L^2(Q), \|F\|_{L^2(Q)}) \cdot \epsilon} \, d\epsilon < \infty,
\]

(19)

where the supremum is over all finitely supported measure \( Q \) on \( \mathcal{X} \). Further if \( Pf^2 < \infty \), then \( F \) is \( P \)-Donsker.

**Sketch of Proof**

Let

\[
\mathcal{F}_\delta := \{(f, g) : f, g \in \mathcal{F}, \|f - g\|_{L^2(P)} \leq \delta\},
\]

(20)

and \( \mathbb{G}_n := \sqrt{n}(P_n - P) \), \( \mathbb{G}_n \in L^\infty(\mathcal{F}) \), i.e.

\[
\mathbb{G}_nf = \sqrt{n}(P_n - P)f = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f(X_i) - \mathbb{E}_P[f(X)]).
\]

(21)

Then

\[
P \left( \sup_{\|f - g\|_{L^2} \leq \delta} \left| \mathbb{G}_n(f - g) \right| \geq \epsilon \right) = P(\|\mathbb{G}_n\|_{\mathcal{F}_\delta} \geq \epsilon) \\
\leq \frac{2}{\epsilon} \mathbb{E} \left[ \sup_{f \in \mathcal{F}_\delta} \left| \sqrt{n}P_n^o f \right| \right] \\
\leq \frac{C}{\epsilon} \mathbb{E} \left[ \int_{0}^{\infty} \sqrt{\log N(\mathcal{F}_\delta, \|\cdot\|_{L^2(P_n)}, \epsilon)} \, d\epsilon \right].
\]

(22)

Let \( \theta_n = \sup_{f \in \mathcal{F}_\delta} |P_n f|^2 \). Note that

\[
N(\mathcal{F}_\delta, L^2(P), \epsilon) \leq N(\mathcal{F}, L^2(P), \epsilon/2)^2,
\]

(23)

we have

\[
\mathbb{E} \left[ \sup_{f \in \mathcal{F}_\delta} \left| \sqrt{n}P_n^o f \right| \right] \leq C \mathbb{E} \left[ \int_{0}^{\theta_n} \sqrt{\log N(\mathcal{F}, L^2(P_n), \epsilon)} \, d\epsilon \right] \\
\leq C \mathbb{E} \left[ \int_{0}^{\infty} (\epsilon \leq \theta_n) \sup_{Q} \sqrt{\log N(\mathcal{F}, L^2(Q), \epsilon)} \, d\epsilon \right] \\
= C \mathbb{E} \left[ \int_{0}^{\infty} 1(\|F\|_{L^2(P_n)} \leq \theta_n) \|F\|_{L^2(P_n)} \cdot \sup_{Q} \sqrt{\log N(\mathcal{F}, L^2(Q), \|F\|_{L^2(P_n)} \epsilon)} \, d\epsilon \right].
\]

(24)

For the remaining steps, we only provide a sketch of the proof. If \( \theta_n \) is small, the dominated convergence theorem implies that the integral goes to 0. If \( \theta_n \to 0 \), applying the Glivenko-Cantelli theorem, we have

\[
\lim_{n \to \infty} \sup_{\|f - g\|_{L^2(P)} \leq \delta} P_n |f - g|^2 \leq O(1) \cdot \delta^2
\]

(25)

with probability 1. Hence if \( \delta \to 0 \), the integral converges to 0. \( \square \)